

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model.

Run on: May 3, 2003, 19:54:31 ; Search time 67.908 Seconds

(without alignments)
13439.802 Million cell updates/sec

Title: US-10-027-000-1

Perfect score: 2976
Sequence: 1 ttatagtcgttgaataa.....aaaaaaaaaaaaaaaaa 2976

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 443362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : Issued_Patents_NA:*

1: /cgn2_6/ptodata/1/lna/5A.COMB.seq:*
2: /cgn2_6/ptodata/1/lna/5B.COMB.seq:*
3: /cgn2_6/ptodata/1/lna/6A.COMB.seq:*
4: /cgn2_6/ptodata/1/lna/6B.COMB.seq:*
5: /cgn2_6/ptodata/1/lna/PTCUS.COMB.seq:*
6: /cgn2_6/ptodata/1/lna/Backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	184.4	6.2	3241	US-09-434-288-11
2	179	6.0	2401	US-09-320-878-20
3	175.2	5.9	2430	US-09-105-537-23
4	175.2	5.9	13613	US-09-105-537-23
5	138.6	4.7	50937	US-09-428-517-1
6	135.8	4.6	2166	US-09-134-078-5
7	111.6	3.8	440365	US-09-103-840A-2
8	111.6	3.8	4411529	US-09-103-840A-1
9	101.4	3.4	16836	US-09-147-236-1
10	101.4	3.4	16836	US-09-147-236-1
11	84.6	2.8	12588	US-08-387-942C-1
12	84.6	2.8	390	US-09-197-649-7
13	83.8	2.8	3033	US-08-462-080B-1
14	83.8	2.8	3033	US-08-462-090-1
15	83.8	2.8	3033	US-08-462-461-1
16	80.4	2.7	1140	US-09-023-173-4
17	78.6	2.6	2064	US-08-343-428-1
18	78.6	2.6	7218	US-08-232-463-14
19	78.2	2.6	1294	US-09-025-691-2
20	78.2	2.6	1521	US-08-083-948-14
21	78.2	2.6	1521	US-08-393-785-14
22	78.2	2.6	1521	US-08-475-694-14
23	78.2	2.6	1521	US-08-712-057-14
24	78.2	2.6	1647	US-08-083-948-13
25	78.2	2.6	1647	US-08-393-785-13
26	78.2	2.6	1647	US-08-475-694-13
27	78.2	2.6	1647	US-08-712-057-13

28	78.2	2.6	1865	US-08-083-948-7	Sequence 7, Appl
29	78.2	2.6	1865	US-08-393-785-7	Sequence 7, Appl
30	78.2	2.6	1865	US-08-475-694-7	Sequence 7, Appl
31	78.2	2.6	1865	US-08-712-057-7	Sequence 7, Appl
32	78.2	2.6	2182	US-09-347-878-35	Sequence 35, Appl
33	76.4	2.6	440365	US-09-103-840A-2	Sequence 2, Appl
34	76.2	2.6	2712	US-09-025-691-4	Sequence 4, Appl
35	74.4	2.5	1143	US-08-387-942C-19	Sequence 19, Appl
36	71.8	2.4	1320	US-08-461-775-8	Sequence 8, Appl
37	71.8	2.4	1320	US-09-031-606-8	Sequence 8, Appl
38	71.8	2.4	1620	US-08-461-775-10	Sequence 10, Appl
39	71.8	2.4	1620	US-09-031-606-10	Sequence 10, Appl
40	70.8	2.4	1929	US-09-380-420C-1	Sequence 1, Appl
41	70.8	2.4	13842	US-09-105-537-30	Sequence 30, Appl
42	70.8	2.4	36778	US-09-105-537-5	Sequence 5, Appl
43	70.8	2.4	38506	US-09-320-878-19	Sequence 19, Appl
44	69.6	2.3	2219	US-08-510-646B-17	Sequence 17, Appl
45	69.2	2.3	2668	US-08-461-775-11	Sequence 11, Appl

ALIGNMENTS

RESULT 1
US-09-434-288-11
; Sequence 11, Application US/09434288
; Patent No. 6303767
; GENERAL INFORMATION:
; APPLICANT: Bellach C., Melanie
; APPLICANT: McDaniel, Robert
; TITLE OF INVENTION: POLYMERASE SYNTHASE ENZYMES AND RECOMBINANT DNA
; FILE REFERENCE: 30062-20030.00
; CURRENT APPLICATION NUMBER: US/09/434,288
; PRIOR FILING DATE: 1999-11-05
; PRIOR APPLICATION NUMBER: 60/107,093
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patent Ver. 2.1
; SEQ ID NO 11
; LENGTH: 3241
; TYPE: DNA
; ORGANISM: Streptomyces narbonneensis
US-09-434-288-11

Query Match 6.2% Score 184.4; DB 4; Length 3241;
Best Local Similarity 48.2%; Pred. NO. 1.3e-28;
Matches 1028; Conservative 0; Mismatches 956; Indels 150; Gaps 12;
DB 544 CGCTGGGTTCCACATTCACCAACCTGCTCGAGAGGAGGAGTGAATGATGGGCAAG 603
1151 CCTGGGCCAGCAGCTTCGAGAGTCCATGCGGAGAGGAGGAGGAGGAGGAGGAGG 1210
DB 604 AGGCATCGCTTAAGAGTGGAGTGTGCTGCGGCCACATTCACATTCAGAGCTCC 663
1211 AGGAGCGGCGGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1270
DB 664 CTTCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 723
1271 CACAGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1330
DB 724 CTTGGGCTTCATTCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 783
1331 CGGTGGCCAGATTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1390
DB 784 TGTGCAATGATAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 843
1391 CGGCAACCAACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1450
DB 844 TCCGTAATTCACACATTCGCTTCAGATTCGTCGAGAGTCCAGGAGGAGGAGGAGGAG 903
1451 TCCGCGAGATCGAGTTCCCGGCGGTTCCAG---GCCTCTCAAGAGCCGCGCGCTCTCT 1507

QY	1444	TCGAGACGCCCCATCTGATACACCTTCGCGGGCGCTACACACACGGTCTCTCCATCTTAGGCG	1503
Db	1124	CTGGGCACACGCCCCACAGTCTGCTCCGGACTCTGGGGGGCGCCACTTGACACCATCAAGGCC	1183
QY	1504	AGCAGTGCCTCCAGGCCACAGCGCGCTCCGGGATCGCTGAGAGGGTCTTCAACGAGCCCC	1563
Db	1184	CGCGGGGGTGGCGGGTGGCCAGGGTGAACGAGACGGGGTGAAGAGACCTTCGGGACGCGAG	1243
QY	1564	CTGGTACCCCTTAACGCCACGACATTTGACGAGCTTTTTCACCAAGACGACATGCAAC	1623
Db	1244	ATTCGGGGGGGAGAACCTCAGC-----CCGGGTTTCAACC	1277
QY	1624	TGGTGACTACTACACACCCCAAGGGGGGACAGACGTGGTACGCCAGATCAAGGACAGT	1683
Db	1278	AGGGCCACCACTCGAGCCGGGGACAAAGCGGGGGCCCTGTACGACGGCAACGGCTGACGTGC	1337
QY	1684	ACACGCGCCAGCAGAGACTGACACTTACAGACTTGCGCTGTGCTGTGCGGACACGCAAAAG	1743
Db	1338	CGCGCGACGGCGAGTACCGCATTCGGCGGTCCGTGCACCGCTGGTTACCGCAAGGTGACAG-	1396
QY	1744	CGTACTAAGACGACAGCTCTGTCTGTGACAAAGCAACGCAACGAAGAGTCCCGGCGATGCT	1803
Db	1397	-----CTCGGGCAAGCANA	1409
QY	1804	TCCTGGCTCCGCAACCCCGAGAGACGGGCCATCTCAATCTCGTCAAGGCAACAGT	1863
Db	1410	CGATCGAGGCGGCTCAGGTCTACGCGCAAGTGAAGCCCGCTCTCCAAGCTGACCAAG	1469
QY	1864	ACAAGTTCAAGATGAGTTCTGCTCGCGCAACCACTACACCCCAAGGCGAGACCATG	1923
Db	1470	GCACGACCAAGCT-----CAGATCTGGGGCTTGGCATGA	1505
QY	1924	TCCCGGCGCACGGCTCCCTCCGCTCGCGGGCTGCAAGGTCAATTGACGACACGAGCCGAA	1983
Db	1506	GTGCGCAACCCGCTCTCCCTGAGACTGAGCTGAGGTGAGTACCGCGCGGGCGGCGAGCGCACGA	1565
QY	1984	TGCAAAAGTCTCGTCCCTTCCCAAGAGACGACCAAGCATCACTGTGGCGCGGCTTAA	2043
Db	1566	TCGCGAAGGGCGTGGAGTGGCGCGGAAGCCCGTACGGCGGTCTGTTTGGC-----	1617
QY	2044	ACGCGCACTGGAGAGCGAGGGGGCGGACCGCGCGAGCATGAAGCTCCCGCGCTGTGG	2103
Db	1618	-CTACGACGACGAGCACCGAAGGGGTGACACCGTCCGAACCTGTGCTGACCGGATACGAGG	1676
QY	2104	ACGAGCTCATTTGCCAGCTGGCGCGCGCGCAACCCAAACACCGTCTGTATGACGAGCG	2163
Db	1677	ACAAGCTGATCTCGCTGTGCGGACGCCAACCGAACAGATGTGTTCTTAACACCG	1736
QY	2164	GCACCCCGGAGAGATGCTCCGTGGCTGACAGCCACAGCCCGCGTCAATCAGAGCGTGTAGC	2223
Db	1737	GTTTGTGGTGTCATGTATGCTCGGTGTGTCAAAGACCCGCGGGTCTGAGCATGTGTAC	1796
QY	2224	GCGGCAACGAGACGGGCACTTCATTGCCGACGTCTTTGGGACTTACAAACCCCTGG	2283
Db	1797	CGGGCCAGAGCGGGGCGGAGGCCACCGCGCGCTGCTACGGGTGAGCTCAACCCGAGCG	1856
QY	2284	GCAAGCTGTCCCTCAAGCTTCCC-----AAGCCCTTGACGAGACAACCCGCGTTTC	2334
Db	1857	GCAAGCTTACCAAGACTTCCCAGGACGAGAACACGAGCGGTGCGCGGCGACCCGA	1916
QY	2335	TCAACTTCCGACCGAGGCGGGGCGACGCTGTACCGGAGAGACGTCTACTTCGGGTACA	2394
Db	1917	CCAGTACCCCGGGGTGCAACACGACGAGAGTACCGGAGGGGATCCAGCTTCTGGGTACC	1976
QY	2395	GGTACTACGAGTTTGGCGACGAAGGACGTCATTTCCTCCCTTGGGCGACGGCTGTCTACA	2454
Db	1977	GCTGGTTGACAAGAGAACTCAAGCCGCTGTTCCGTTGGGGACAGGCGCTGTCTACA	2036
QY	2455	CCACTTTTGCCTTTTCAATCTCTCCGTGCTCACAGAGAC---GGCAGGTGAGCGTGT	2511
Db	2037	CTGTGTTACGCAAGAGCGCCCGACCGTGTGCTGACTTCACAGGGTGTCTGAAGGTCA	2096
QY	2512	CCCTTCCGTGAAGAACACCGGCTCCGTCCGGCGGCGACAGGTGGCCAGTCTACGTC	2571

Db 2097 CGGTACGGTCCGACACACCGGAGCGCGCCAGAGAGCTCTCCAGGCGTACCTCG 2156
QY 2572 AGCCCTTCAGACGGGCAAGATTAAACCGCCCGCCAGAGGCTCAAGGCTTCGCAAG 2631
Db 2157 GTGCCAGCCCGAAGCTGAGCGCTCCGCGAGCGCAAGAAAGAGCTCTGGGCTACAGGAAG 2216
QY 2632 TCGAAGTGCAGCCCGCGAGAGCAAGGCGGTGAC 2665
Db 2217 TCTCGCTCCGCGCGGCGGAGCGGAGAGCGGTGAC 2250

RESULT 3

US-09-105-537-23
; Sequence 23, Application US/09105537A
; Patent No. 6265202
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600.438051
; CURRENT APPLICATION NUMBER: US/09/105,537A
; CURRENT FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 23
; LENGTH: 2430
; TYPE: DNA
; ORGANISM: Streptomyces venezuelae
US-09-105-537-23

Query Match 5.98; Score 175.2; DB 4; Length 2430;
Best Local Similarity 47.68; Pred. No. 8.8e-27;
Matches 1016; Conservative 0; Mismatches 968; Indels 150; Gaps 11;

QY 544 CGCTCGGTTCCACATTCATCAACCACTGCTCGAAGAGGCAAGTATGATGAGGCAAG 603
Db 353 CCCTGGCCAGACCTTGGAGCAGCACATGGCCGACAGCTAGGCAAGTATGAGGCGCG 412
QY 604 AGCCATCGCTAAGAGTGGCGATGATGCTCGGCGCCGACTATCAACAGCTCC 663
Db 413 AGGTCGCGGCTCAACACGAGCATGGTCTGGGCGCGATGATGAACAACATCGGCTGC 472
QY 664 CTCTCGGAGAGCGGCTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 723
Db 473 CGCACGGGCGGCACTAGAGACCTTCAAGCGAGGAGCGCCCTGCTCTCCGCGACCG 532
QY 724 CTGGGCTCTCATCCGCGCATTCAGAGCATGAGTGCAGGCTAGCATCAACACTTTT 783
Db 533 CGGTGCGCGAGATCAAGAGGATCCAGGGTGTGATGATGATGATGATGATGATGATGAT 592
QY 784 TGTCCATGATCAAGAGGAGAGCGATGATGATGATGATGATGATGATGATGATGATGAT 843
Db 593 CGGCAACAACACGAGAAACACCGCTTCTCGTGAACGCAATGTCGACGACAGACGG 652
QY 844 TCCGTGAATCTAGCACTCCGTTCCAGATGCTGTTGGAGAGTCCCAAGCGCGGCTGCT 903
Db 653 TCCGCGAGATCGATTCGCGGCTTGCAG---GCTCTTCAAGGCGCGGCGCTCTCT 709
QY 904 TCAATGAGCGGTACATGAGCATGAGCGTGTCTGTCAGACGAGAAACCTTAATATCTTG 963
Db 710 TCAATGAGCGGTACATGAGCATGAGCGTGTCTGTCAGACGAGAAACCTTAATATCTTG 769
QY 964 ATGAGATGCTTCAAGAGATGAGGTTGGATGAGCTTAATCATGAGCGAGTGTGAGCA 1023
Db 770 ACAAGGCTGCGCAGCGAGTGGGCTTCCAGGGCTGGGATGATGATGATGATGATGATGAT 829
QY 1024 CATACAGTACAGAGAGCGTGTGTCAGAGCGCTGACCTCGACCTGCGGAGCTCCAC 1083
Db 830 C---CCGCGGACACGAGCGCATTCACAGGCGCTCGACAGAGATG----- 873

QY 1084 GCTTCGAGAGAAACATCAAGTTCAGAGCTTCCAGAGGAAACCTTTATTCACGTCA 1143
Db 874 -----GGGTGAGACTCCCGGCGAGCTCCGAAAGGCGAGCCCTCCCGCGGCA 925
QY 1144 TTGACAGAGGAGGTTAGGAAAGTCTTCAAGTTCAGAGAGTGTGCTCCGAGTGA 1203
Db 926 AGTCTTCGCGGA---GGCTGAAGAGCGCGCTCTGAAGGAGCAGGTCTCCGAGCGG 982
QY 1204 CGGAGAGCGCCCGAGAGCAGTGTCAACAACACCCCGCAAGCGAGCTCTCTCCGA 1263
Db 983 CGGTGAGCGGTGCGGAGAGGATGTCGCGCAAGATGAGAGATGTCGCTCTCTCTG 1042
QY 1264 AGTTGGCAACAGAGGATGCTGCTGTCAGAGCAAGCAAGCAAGTTCGCTTGAAGA 1323
Db 1043 CCACTCCGCGCGCGCGCGCGAGAGCAAGCGGAGTCCGAGGAGTGTCCCGAAG 1102
QY 1324 AGAAGAAAGAGCGCTGATTTGCGCCCAAGCGCAAGCGCAGTATACAGCGCGAG 1383
Db 1103 TCGCGGAGAAAGCGCGGTCTCTCGCAACAGAGGCGCGCTGCGCTGCGCGGTG 1162
QY 1384 GCTTCGCGCAGTCAAGGCGCTTCAAGCAAGTCACTCTTTCAGCAAGAGCATGACG 1443
Db 1163 AGCGCGCAAGAGCATTCGCGGTATGCGCCCGAGCGCCCTCGAC-CCAAAGTCAAGCG 1221
QY 1444 TCGAGAGCGCGCATGATGACCGGTGCGGCTTACACACACCGTCTCTCCATTAGCG 1503
Db 1222 CTGGGAGAGCGCCACAGTCTGCGGACTGCGGCGCGCGCGCACTGACACATCAAGGCG 1281
QY 1504 AGCAGTGCCTCAAGCGCGCGAGCGGCTCGCGGATGCGGTGAGAGGCTTCAACAGAGCC 1563
Db 1282 CGCGCGGCTGCGGCTGCGGAGTACAGAGAGCGGTGAGAGAGCTTCGCGAGCGAG 1341
QY 1564 CTGTTACCCCTTAAACCGCGACATGACAGCTCTTTCACCAAGAGCATGACGAC 1623
Db 1342 ATCCGCGGCGGAGACCTCAGC-----CGCGCTTCAAGC 1375
QY 1624 TGTGAGTACTACTACACCCCAAGAGCGGAGACAGTGTGACGAGTGAAGGACAGT 1683
Db 1376 AGGCGCACAGTCTGAGCGCGGCAAGGCGGCGGCTGATGACAGGAGCGGTGAC 1435
QY 1684 ACACCGCGCGAGAGGATGACCTACAGTGAAGCTGGGCTGTGCGCGC-AGGAAAG 1743
Db 1436 CGCGCGAGCGAGTACGCGATGCGGCTGCTGCGCAGCGGTGATGACAGGCTGACG 1494
QY 1744 CGTAGTAGACAGCAGCTGCTGTGACAAACGCCAACAGAGGTCCCGAGTGTCT 1803
Db 1495 -----CTCGGAGCGCA 1507
QY 1804 TCTGCGCTCCGCGCACCGCGAGAGAGCGGCGCGATCATCTGTCAGAGGCAACAGT 1863
Db 1508 CCAAGAGCGCGGTGAGTCTTACGGAAGGTGAGAGCGCCCTCTCTCAAGCTGACAAAG 1567
QY 1864 ACAAGTCAAGATCGAGTTCGCTCGGACCCACTTACACCTCAAGGCGAGCAACATCG 1923
Db 1568 GCAGCACAAGCT-----CACATCTCGGCTTCCGATGA 1603
QY 1924 TCCCGGCGCAGGCTCTCCGCTCGCGGCTGCAAGGTATGACAGCAGCGGAGAA 1983
Db 1604 GTGCCACCCGCTCTCTGAGAGTGGGCTGAGTACCGCGCGGCGCGAGCAGCA 1663
QY 1984 TCGAAGATCGGTGCGCTCCGCGCAAGGAGCAGCAAGGTATCATGTCGCGGCGCTTA 2043
Db 1664 TCGGAGAGCGGTGAGTTCGCGGAGAGGCGCGTACGCGGCTGTCTTGC----- 1715
QY 2044 AGCGGACTGGAGAGCAGAGGCGCGAGCGCGAGCATGAGCTCCCGCGCTGCTGG 2103
Db 1716 -CTAGAGAGCGAGCAGAGGCGGTGAGCGGTGCGAAGCTGTGCTCCGAGTACGAGG 1774
QY 2104 ACAAGCTATGTCGAGTGTGCGCGCGGAGAACCAACACCGTGTGTCATGACAGCG 2163
Db 1775 ACAAGCTATGTCGAGTGTGCGCGAGCAGCAACCGAGACAGATGCTGCTCAACACCG 1834
QY 2164 GCACCCCGAGAGATGCGCTGCTGACGCGACGCGCGCGGTATCCAGCGCTGTGAG 2223

Db 1835 GTTCTGCGGTGATGATGCGGTGCTGTCCAGACCCGCGGTCTCGGACATGTGTAC 1894
QY 2224 GCGGCAACGAGAGGCGCACTCATTTGCGAGCTCAACCCCTCGG 2283
Db 1895 CCGGCGAGGCGGCGCGAGGCGCGCGCTGTCTACGCTGACGTCACCCGAGG 1954
QY 2284 GCAAGGTGTCCTCAGCTTCCG-----AAGCGCTGACAGCAACCCCGCTTC 2334
Db 1955 GCAAGGTGTCCTCAGCTTCCGCGCGCGAGCAACGAGCGGTGCGCGGAGCCCA 2014
QY 2335 TCAACTTCCGACGAGGCGCGGCGACGCTGACGCGAGGAGCGCTACGCGGTGTA 2394
Db 2015 CAAAGTACCGCGGCGGTGACCAACGACGACGACGCGGAGGAGCTACGCGGTAC 2074
QY 2395 GGTACTACGAGTTCGCGCAACGAGCTCAATTTCCCTTTGGCGACGCGCTGTCTACA 2454
Db 2075 GCTGTGTCGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2134
QY 2455 CCACTTTGCTTTCCATCTCTCGGTGTCTACAGAGC---GGCAGCTAGCGTGT 2511
Db 2135 CTTGTCAGGAGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2194
QY 2512 CCGCTTCCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2571
Db 2195 CCGTACGAGTCCGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2254
QY 2572 AGCCCTTCGAGAGGCGCGCAATTAACCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2631
Db 2255 GTGCGACCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2314
QY 2632 TCGAATGTCAGCCCGCGGAGAGCAGAGGCGGTGAC 2655
Db 2315 TCTGCTCGCGCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2348

RESULT 4
US-09-105-537-3

Sequence 3, Application US/09105537A
Patent No. 6265202
GENERAL INFORMATION:
APPLICANT: Sherman, D.H.
APPLICANT: Liu, H.
APPLICANT: Xue, Y.
APPLICANT: Zhao, Y.
TITLE OF INVENTION: DNA encoding methymycin and pikromycin
FILE REFERENCE: 600.438051
CURRENT APPLICATION NUMBER: US/09/105,537A
CURRENT FILING DATE: 1998-06-26
NUMBER OF SEQ ID NOS: 43
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 3
LENGTH: 13613
TYPE: DNA
ORGANISM: Streptomyces venezuelae
US-09-105-537-3

Query Match 5 98: Score 175.2; DB 4; Length 13613;
Best Local Similarity 47.68; Pred. NO. 1.5e-26;
Matches 1016; Conservative 0; Mismatches 968; Indels 150; Gaps 11;

QY 544 GCGTGGTTCACATTAACCAACTGTCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 603
Db 4664 CCTTGCGACGACCTTTCAGACGACGACGACGACGACGACGACGACGACGACGACGACG 4723
QY 604 AGGCGATCGTAAGAGTGGCGAGTGTGCTGCGCGCGACGACTACGACGACGACGACG 663
Db 4724 AGGCGTGGCGTCAACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4783
QY 664 CTTGCGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 723
Db 4784 CCGACGCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4843

QY 724 CTTGCGTTCATCCGCGGACATTCAGAGCAGTGGAGTGCAGAGCTACGATCAAGCACTTTT 783
Db 4844 CCGTGGCCCAATGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4903
QY 784 TGTGCAATGATCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 843
Db 4904 CCGGCAACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4963
QY 844 TCCGGAATCTACGACCTCCCGGTCGAGATTTGCTGCGAGAGCTCCGAGCGGAGTGGCT 903
Db 4964 TCCGCGAGATGAGTTCGCGCGGTCGAG---GCGTCTCCAGAGCGCGCGCGCTCTCT 5020
QY 904 TCATGACGCGTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 963
Db 5021 TCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5080
QY 964 ATGGATGCTTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1023
Db 5081 ACAAGGTGCTGCGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 5140
QY 1024 CATACAGTACACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1083
Db 5141 C---CCGCGGACCGAGCGCATCACCAGGCGCTGACAGGAGATG----- 5184
QY 1084 GCTTCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1143
Db 5185 -----GGCGTGCAGCTCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 5236
QY 1144 TTGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1203
Db 5237 AGTTCTTTCGCGCA---GGCGTGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 5293
QY 1204 CCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1263
Db 5294 CCGTACGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 5353
QY 1264 AGGTTGCAACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1323
Db 5354 CCACTCCG 5413
QY 1324 AGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1383
Db 5414 TCGCCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 5473
QY 1384 GCTTGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1443
Db 5474 ACCCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 5532
QY 1444 TCGAGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1503
Db 5533 CTGGGAGGCG 5592
QY 1504 AGCAGTGCCTACG 1563
Db 5593 CCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 5652
QY 1564 CTGGTACCGCTAACCG 1623
Db 5653 ATCCCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 5686
QY 1624 TGTGTGACTACTACCG 1683
Db 5687 AGGCGACACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 5746
QY 1684 ACACCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1743
Db 5747 CCGCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 5805
QY 1744 CGTACGTAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1803
Db 5806 -----CTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 5818
QY 1804 TCTTGGGCTCGGACCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1863

```

Db 5819 CCATGAGGCGCGTCAAGTCTACGGCAGGTGAGCAGCCCGCTCTCAAGCTGACCAAG 5878
QY 1864 ACAATTCAAGATTCAGATTGCGCTCCGACCCACTACACCTCAAGGGCCACACATCG 1923
Db 5879 GCAGCACAAGCT-----CAGATCTCGGGCTTCGGAGTA 5914
QY 1924 TCCCGGCGCAGCGCTCCCTCCGCGCTGCGCGCTGCAGATTCATTCAGCAGCCGAA 1983
Db 5915 GTGCCACCCCGCTCCCTCGAGTGGGTGAGTACCCCGCGCGCCACCGACGA 5974
QY 1984 TCGAAATGCTGCTCCCTCCGCAAGGACACAGCAGTATCATCTGCGGGGCTTA 2043
Db 5975 TCGCAGGCGCGTGAATGCGCGCGGAGGCCGCTACCGCGCGCTGCTTCG----- 6026
QY 2044 AGCGGAGTGGGAGACCGGAGCGCGCGCGAGCATGACCTCCCGCGCTGCTG 2103
Db 6027 -CTAGCAGAGGAGCGACCGAGGCGCTGACCGCTCCGAACTCTGCTGCGGGTACGAG 6085
QY 2104 ACCAGCTCATTCGCGACGTGCGCGCGCGGAAACCAACCGCTGCTCATGCGAGCG 2163
Db 6086 ACAAGCTGATCTCGGCTGTCGCGAGCGCAACCGGAACAGATGCTGCTCAACACG 6145
QY 2164 GCACCCCGGAGAGATGCTGCTGCTGCGACGCGACCGCGCTGATCAGGCGCTGAGC 2223
Db 6146 GTTGCTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6205
QY 2224 GCGGCAAGAGAGCGGCAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2283
Db 6206 CGGGCGAGGCGGGGCGCGAGCGCGCGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTG 6265
QY 2284 GCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2334
Db 6266 GCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2325
QY 2335 TCACTTCCGAGCGGAGCGGCGCGGCGCGAGCTGACGCGGAGGAGCTGCTGCTGCTG 2394
Db 6326 CAACCTACCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6385
QY 2395 GGTACTGAGATTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2454
Db 6386 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6445
QY 2455 CCACTTTCGCTTTCATCTCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2511
Db 6446 CTTGCTTACGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 6505
QY 2512 CCTCTCCGTAAGAACACCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2571
Db 6506 CGGTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6565
QY 2572 AGCCCTCCGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2631
Db 6566 GTGCCAGCGCGGAGAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 6625
QY 2632 TCGAAGTGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2665
Db 6626 TCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6659

```

```

RESULT 5
US-09-428-517-1/c
; Sequence 1, Application us/09428517
; Patent No. 6251636
; GENERAL INFORMATION:
; APPLICANT: Belilach, Mary C.
; APPLICANT: Shah, Sanjay Krishnakant
; APPLICANT: McDaniel, Robert
; APPLICANT: Tang, Li
; TITLE OF INVENTION: RECOMBINANT OLEANDROLE POLYKETIDE SYNTHASE
; FILE REFERENCE: 30062-20029.00
; CURRENT APPLICATION NUMBER: US/09/428, 517
; CURRENT FILING DATE: 1999-10-28

```

```

; EARLIER APPLICATION NUMBER: 60/120,254
; EARLIER FILING DATE: 1999-02-16
; EARLIER APPLICATION NUMBER: 60/106,100
; EARLIER FILING DATE: 1998-10-29
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patent Ver. 2.1
; SEQ ID NO 1
; LENGTH: 50937
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Recombinant DNA
US-09-428-517-1

```

```

Query Match 4.7%; Score 138.6; DB 4; Length 50937;
Best Local Similarity 54.0%; Pred. No. 6,4e-19;
Matches 354; Conservative 0; Mismatches 289; Indels 12; Gaps 3;

```

```

QY 2048 CGACTGGAGAGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2107
Db 3417 CGAGCAGCGCTCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3358
QY 2108 GCTCATTCGCGAGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2167
Db 3357 CTTGATCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3298
QY 2168 CCGCGAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2227
Db 3297 CTCTCTACCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3238
QY 2228 CAACGAGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2287
Db 3237 CGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3178
QY 2288 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2341
Db 3177 GCTGACCGCGAGACCTTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3118
QY 2342 CCGCAGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2398
Db 3117 CTACCGGAGGAGTGCATGACCACTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3058
QY 2399 CTACGAGTTGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2458
Db 3057 GTACGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2998
QY 2459 TTTTTCCTTTCATCTCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2515
Db 2997 CTTGACTTACCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2938
QY 2516 CTCTGTAAGAACACCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2575
Db 2937 CACGCTGCGCAACCGGCTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2878
QY 2576 CTTCAACGCGCGCAAGATTAAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2635
Db 2877 GTCCCGGACGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2818
QY 2636 ACTGCGCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2690
Db 2817 GCTGCGCGCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2763

```

```

RESULT 6
US-09-134-078-5
; Sequence 5, Application us/09134078
; Patent No. 6368844
; GENERAL INFORMATION:
; APPLICANT: Byline, Edward J.
; TITLE OF INVENTION: GLYCOSIDASE ENZYMES
; CORRESPONDENCE ADDRESS: 72
; ADDRESSEE: Gray Cary Ware & Freidenrich LLP

```


1

Db	12965	ATCCGGTCGACGACCGGGCGCATGTGTGGGCGACCGACACCATTTGACGGCGGTGCAGTGCAGACATG	13024
QY	760	TGCAGGCTACGATCAGACGACTTTTGTGTCAATGATCTCAGGAGGACAGCGCATGATGTGTC	819
Db	13025	TGATCTCCAGCGCTCAAGCACTATTATGGATGAGATGACCTCGAAACCTCGCGCATGACCATTA	13084
QY	820	AGAGCANTCGCAGCAGGAGCGGGCTCGGTGAAATCTACGCACTCCGCTTCACAGATTGCTG	879
Db	13085	GCGCGGATATCGACCTGTGTGGCCATGCTGTGAAGACGACGTGTGTGGGTTGTGAAATGGCC	13144
QY	880	TGCGAGACTCCACGCGGGGTGCGTTTCATGACGGCGTACATGCACTCATATGGCTGTCT	939
Db	13145	TTGAAACCGGGGACCTCGGGCGGTGTCATGTGCTGTACAAACCGGTGTAAAGACCTTTATG	13204
QY	940	GCAACCGAACCCTTAATATCTTGATGGATGCTCCGAAAGCAATGGGGTTTGGATGAGGC	999
Db	13205	CGTGTGAAACCGCTGTACTCTGTGAAACAGACCTTAAAGCAGAACTGTGGCATTATCCGGGCT	13264
QY	1000	TAACTATGACGCACTGGTAGCGACATACATGTAACCAAGAAACCGTTGTGGCAGGCTCG	1059
Db	13265	TTGTTCATGTCCGACTGGGGGGCCACGACCATTTCTCCGCGCGGGCGGCGCTGGCGGGGTGG	13324
QY	1060	ACCTCGA	1066
Db	13325	ATCAGGA	13331

RESULT 10
US-09-147-236-10
; Sequence 10, Application US/09147236A

```

: GENERAL INFORMATION:
: APPLICANT: TONOUCHI, Naoto
: APPLICANT: TSUCHIDA, Takayasu
: APPLICANT: YOSHINAGA, Fumihito
: APPLICANT: TAHARA, Naoki
: APPLICANT: HAYASHI, Takahisa
: TITLE OF INVENTION: NOVEL GENE, GROUP OF GENES, AND NOVEL BETA-GLUCOSIDASE
: FILE REFERENCE: 6537-011-0PCF
: CURRENT APPLICATION NUMBER: US/09/147,236A
: CURRENT FILING DATE: 1999-04-08
: EARLIER APPLICATION NUMBER: PCT/JP97/03633
: EARLIER FILING DATE: 1997-10-09
: NUMBER OF SEQ ID NOS: 12
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 10
: LENGTH: 16836
: TYPE: DNA
: ORGANISM: Acetobacter xylinum
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (1891)..(2922)
: FEATURE:
: OTHER INFORMATION: Nucleotide sequence is the same as SEQ ID NO:1
: FEATURE:
: OTHER INFORMATION: n at positions 15741 and 15767 may be a, g, c, or
: OTHER INFORMATION: t
: US-09-147-236-10

```

Query Match	3.4%	Score 101.4	DB 4	Length 16836
Best Local Similarity	50.5%	Pred. No. 1.6e-11		
Matches 246	Conservative 0	Mismatches 241	Indels 0	Gaps 0

Accession	Sequence	Position
QY 580	AGGCAAGCATATGGGCAAAAGAGGCCATCCCTAAGATGGCGCATATGGATCTCTGGCC	639
Db 12845	AGGGCGGTGCATGATCGGGCGCGAGGCATGGCAGAGCGGCTTCAACATCTCTTGGCG	12904
QY 640	CGACTATCAACATGCAACGCTCCCTCGTGTGACGTGCTCGATGCATGTGTGAG	699
Db 12905	GGGGTGGGACCTACGCGCGACGCCGCTGTGGCGGCCGCACATTTGAAATATGGCGGGGAG	12964
QY 700	ATCCGTCCTGGCGGCTTGGAAGCTCGGCTCTCATCCGGGCGCATTCAGAGCACTGGAG	759

Db	12365	ATCCGCTGCAGACCGGGCCGATGCTGGCCAGCACCACTTTGCAGGCGTGTGCACTGTCAGACATG	13024
QY	760	TGCAGGCTACGATCAGACACTTTTGTGTCAAATGATATCAGAGAGACAGCGCATGATGCTGC	819
Db	13025	TGATCTTCACGCTCAACGACTATATGGATGATGTAACCTCGAACCCTGCGCATCACCATCA	13084
QY	820	AGACATCTGTCAACGAGCGGGCTCTCCGTGAATCTACGACACTCCGTTTCCAATATGTCTG	879
Db	13085	GCGGGGATATTCGACCCTGTGGCCATGCGCAAGAAAGGACCTGCTGGCTTTGGAATGGCGC	13144
QY	880	TGCAGACTCCCAAGCCGGGTGGCTTCATATGACGGCGCTACAAATGGCATCAATGGGCTGTCGT	939
Db	13145	TTGAAACCGGGGACTCCGGGGCGGGTATATGTGCTGTAAACCGGCTAAAGACCTGTATG	13204
QY	940	GCACGAGAAACCTTAATATCTTGATGGGATCTTCGAAAGGAATGGGTTTGGCATGGCC	999
Db	13205	CGTGTGAAAAACCGGTACTGCTCTTAACAAACCGCTTAACGASGACTTGGCATTTATCTGGGCT	13264
QY	1000	TAAATCATGAGAGACTGTGATGCGACATATACATACCAAGAAACCGTTGTGGCAGGCGCTG	1059
Db	13265	TTGTGATGTCCGACTGTGGGGGCCACGACCATTTCTCCGGCGGGCGCGGTGGCGGGGCTGG	13324
QY	1060	ACCTGCA	1066
Db	13325	ATCAGGA	13331

RESULT 11
US-08-387-942C-1

GENERAL INFORMATION:
APPLICANT: ERTESVAG, HELGA
APPLICANT: VALLA, SVEIN
APPLICANT: SKJAK-BRAEK, GUDMUND
APPLICANT: LARSEN, BJORN
TITLE OF INVENTION: DNA COMPOUNDS COMPRISING SEQUENCES
TITLE OF INVENTION: ENCODING MANNURONAN C-5-EPIMERASE
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH, LLP
STREET: P.O. BOX 747
CITY: FALLS CHURCH
STATE: VA
COUNTRY: USA
ZIP: 22042
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/387,942C
FILING DATE: 09-MAY-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: MURPHY JR, GERALD M.
REGISTRATION NUMBER: 28,977
REFERENCE/DOCKET NUMBER: 1809-106P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-205-8000
TELEFAX: 703-205-8050
INFORMATION FOR SEQ. ID NO.: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 12588 base pairs
type: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: Acetobacter vinelandii
STRAIN: E

```

?      FEATURE:
?      NAME/KEY: CDS
?      LOCATION: 290..1951
?      FEATURE:
?      NAME/KEY: CDS
?      LOCATION: 2227..6438
?      FEATURE:
?      NAME/KEY: CDS
?      LOCATION: 6702..9695
?      FEATURE:
?      NAME/KEY: CDS
?      LOCATION: 9973..12588
?      US-08-387-942C-1

```

Query Match	Best Local Similarity	Score	DB	Length
2.8%	43.3%	84.6	2	12588
Pred.		3.8e-08		
Matches	451	Conservative	0	Mismatches 584
				Indels 6
				Gaps 1
QY	1535	CATGCGCTGAGGGGTTCTTCAACGAGGCCGCCCTGGTATACCCTCTAACCCGACAGCATTTGACGA	1594	
Db	2601	CATCCCCCGGCGAGAGGGCCCGCATGCGCGACGTCGACCTCGGAGCGGGGTGGAAATCCGCGCA	2660	
QY	1595	GCTCTTCTTACCAACAGCGGACATGACCTGGTGGACTACTTACCAACCCCAAGCGGACGA	1654	
Db	2661	GATGTCCGAGCTACGTTTTCAGCCGCCACGACGAGACATCAACCTTACATCCGCGACAG	2720	
QY	1655	CACGTGGTACCCGACATGAGAGGGACGTAACCCGCGACGAGAGCATGTGACCTACAGAGCT	1714	
Db	2721	CGTGGCCCGACGAAACAGCGCTCGACGGCTTCGTCGGCGACTACACAGTGGGGGTGTT	2780	
QY	1715	CGGCTCTGTCGTGTGGGGGACGCGCAAAAGCGTACGTACGACGACACGCTGTCGTGCACA	1774	
Db	2781	CGAGAAACAAGTCTCTTACCAACACGACGCCACGGCTTCAACATCTGTCACGACACCA	2840	
QY	1775	CGGCACCAAGCAGGTCCCGGGGATGCTCTTCTTGGCTCGCGCACCGCGGAGAGACGGG	1834	
Db	2841	CGACTTCTGTCCTGAGCAACAGTCGGCTTACGCGCAACGGCGGGCGGCTGTGTGTGA	2900	
QY	1835	CGGCATCATCTCTGTCACAGGGCAACACGTAACGTTCAAGATCGAGTTCGGCTCGCAC	1894	
Db	2901	GGCGGGCTGTACGACCTGCGCCCTATCCCTACGACATCTCGATCGACGGGGCGGCTTACTA	2960	
QY	1895	CACCTACACCTCTCAAGGGGACACCATCGTCCCGCGGCGACGGCTCCCTCGCGGTGGGGG	1954	
Db	2961	CGACAACGCTTTGGAAGGGGTGCTGACTCAAGATGGGCCACGACGTACACCTTGACAGACGC	3020	
QY	1955	CTGCAAGGTGATGACGACGACGCGGAAATCGAAATGTAAGTCCGTGCGCTCGCCACAGACA	2014	
Db	3021	CGAGATCTACGCGCAACGGCGCTGTACGGGGTGGCGCTTACGCGGCCCGACGAGTGCAGAT	3080	
QY	2015	CGACCAAGGTATCTGTGGCGGGGCGCTTAAACGCGCAGCTGGGAGACCGGAGCGCGACCG	2074	
Db	3081	CCTGCACAACCGATATCCACGACAAATTCGCACAGACGGCCCTATGTGCGGAATCTGTGTGCA	3140	
QY	2075	CGCGAGCATGAAGCTCCCGGGCGGTGTGAGACAGTCAATTCGCGACGTCGCGCGCGCGAA	2134	
Db	3141	GTCTACGACACACACCGCGGGGTGTCCGGCAACTTTAGCTACACACCGGCACTGGCT	3200	
QY	2135	CCCAAAACCGTCTGTCGATGACGAGCGGGACCCCGGAGAGATGCCCTGTGTCGACGC	2194	
Db	3201	CGAAGGCAACTCATACGCGGCTCGGCCAATTTCCACTTACCGCATTCACGAGGCGCGCGGA	3260	
QY	2195	CACGCCCGCTCATTCACAGGCGCTGTATACGGGCGGCAACGAGACGGGCAACTCCATTGGCGA	2254	
Db	3261	CGCGACCGACTACAGAGCCCTTACGCCAACAGATGACGAGTGTGCAGACCGGGGGGT	3320	
QY	2255	CGTGTCTTTTGGGCACTACAAACCCCTGGGGCAGCTGTCCCTCACTTCCCAAGCGCT	2314	
Db	3321	ACGGCTGTATGGCCCAACTGACGGTTTTCAGCCACTCGGCGAGTGGCCAGACGGCGAC	3380	
QY	2315	GCAGCAACAACCCCGTTCCTTCA-----CTTCCGACCGAGGCGGGCGACGCTGTGA	2368	
Db	3381	CCTGGAAGGACGCGCGGCGCAACGATGCGCTGAGAGGGGACGAGGCCCCACGACGAGCTGCT	3440	

OY	2369	CGGGAGGAGCGCTACAGCTCCGGGTACAGAGTACTACAGAGTTGGCCGACAGAGAGCTGAATTT	2428
Db	3441	CGGGCAGGCGGGCGAGCAGCCGCTGAAGGGCGATGCCGGCAAGCAATCTTCGACGGCGG	3500
OY	2429	CGCCTTTGGCCACGGCCTGTCTCAACACATTTTGGCTTTTCCAAATCTCCGGTGTCA	2488
Db	3501	GGCAGGGGGGAGCAACCTGACCGGGGGCGGGCGCCGACACTTCCGCTTCTCAGCGG	3560
OY	2489	CAAGGACGGCAAGCTAGAGCTGTCCCTTTCCTGGAAGAACACGGGCTCCGATCCCGGCGC	2548
Db	3561	CACGCAAGCTACCGCACCCACAGAGCCGACGTTCAACGACCTGATACCACCCACTTTGAGGC	3620
OY	2549	ACAGGTGGCCAGGCTCAAGT	2569
Db	3621	CGACGAGGACAGCATGACCT	3641

```

RESULT 12
US-09-197-649-7
: Sequence 7, Application US/09197649
: Patent No. 6194550
GENERAL INFORMATION:
APPLICANT: Gold, Larry
APPLICANT: Tuerk, Craig
APPLICANT: Pridnow, David
APPLICANT: Smith, Jonathan D.
TITLE OF INVENTION: Systematic Polypeptide Evolution by Reverse Translation
FILE REFERENCE: NMX02/CI-CON
CURRENT APPLICATION NUMBER: US/09/197,649
CURRENT FILING DATE: 1998-11-23
EARLIER APPLICATION NUMBER: 07/829,461
EARLIER FILING DATE: 1992-01-31
EARLIER APPLICATION NUMBER: 07/739,055
EARLIER FILING DATE: 1991-08-01
EARLIER APPLICATION NUMBER: 07/561,968
EARLIER FILING DATE: 1990-08-02
NUMBER OF SEQ ID NOS: 26
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 7
: LENGTH: 390
: TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Sequence
OTHER INFORMATION: having a 120 repeat of ACG flanked by fixed
US-09-197-649-7
OTHER INFORMATION: fragments having NcoI restriction sites.

```

[illegible]

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/625,140
 FILING DATE: 10-DEC-1990
 ATTORNEY/AGENT INFORMATION:
 NAME: Dillahunt, T. Gene
 REGISTRATION NUMBER: 25,423
 REFERENCE/DOCKET NUMBER: 010055-056
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 415-854-7400
 TELEFAX: 415-854-8275
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 3033 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 HYPOTHEetical: NO
 ANTI-SENSE: NO
 ORIGINAL SOURCE:
 ORGANISM: Trichoderma reesei
 FEATURE:
 NAME/KEY: CDS
 LOCATION: join(311..375, 446..2205, 2270..2679)
 FEATURE:
 NAME/KEY: Intron
 LOCATION: 376..445
 FEATURE:
 NAME/KEY: Intron
 LOCATION: 2206..2269
 US-08-462-090-1

Query Match

Best Local Similarity 2.8%; Score 83.8; DB 3; Length 3033;
 Matches 264; Conservative 0; Mismatches 257; Indels 6; Gaps 2;

QY 552 TCACATTCAACCAACTCTGCTCGAAGAGGCGAGTAAAGTATGGGCAAGAGCCATC 611
 DB 720 TCGACGTGGGATGTCATTTGATCCCGAAGCGAGACATTCATCGTGAAGAGTGAAG 779
 QY 612 GCTAAGAGTGGCATGTGATCTCGGCC--GACTATACATGCAACGCTCCCTCTC 668
 DB 780 GCTCGGGGATTCATGTCATCTGCTGCTGGGCGCGTGGAAAGACTCCGAG 839
 QY 669 GGTGACGTGGCTTCGAGTGGATGGTGGAGATCCCTTCCTGGCGGCTTGGAGCTCG 728
 DB 840 GGGGTGCGCACTGGAGGGCTTCGCTGTCATTCATTCATTCGCGGATTCGATGGGT 899
 QY 729 GCTCTCATCCGCGCATTCAGAGCACTGAGTGCAGGCTACGATCAACACATTTTGTGC 788
 DB 900 CAACCATCAACGCGCATTCAGTGGTGAAGCGTGCAGGCGAGACACATCATATCTCTC 959
 QY 789 AATGATCAGAGAGACAGGCGCATGTGTCAGACATGTCAGGAGCGGCTCTCCGT 848
 DB 960 AAGGACAGAGAGCTCATGACCAACCATTTTGCAGCAACCCAGATCCGAATCTCCAT 1019
 QY 849 GAATCTACGCACTCCGCTTCAGATTGCTGCGAGACTCCGCGGCGGCGGTTCATG 908
 DB 1020 GACCTGTATACCTTGGCCATTTGCCAGCGGTTC--AGCCATATGCTCTTGTATG 1076
 QY 909 ACGGCTTAATGTCATGAGCTGTGTCGACGGAACCTTAATATATTTGATGG 968
 DB 1077 TGGTCTACACAGAGTCAATACCACTGGGCTCGAGGATGATACACGCTCACT 1136
 QY 969 ATGCTCGAAGAGATGGGTTGGATGGCTTAATCATAGACGCTGTACGCAATAC 1028
 DB 1137 GTCTGTAAGACCACTGGGTTGCCAGGCTATGTCATGAGGAGTGAACGACACAG 1196
 QY 1029 AGTACCAAGAGCCGTTGGAGGCTCGACCTCGAGATGCCGG 1075
 DB 1197 ACGAGTGCACAAAGCGCAATTCGGGCTTACATGCAATGCTCG 1243

RESULT 15

US-08-463-461-1
 Sequence 1, Application US/08463461
 Patent No. 6103464

GENERAL INFORMATION:
 APPLICANT: Fowler, Timothy
 APPLICANT: Barnett, Christopher C.
 APPLICANT: Shoemaker, Sharon
 TITLE OF INVENTION: Saccharification of Cellulose by Cloning
 TITLE OF INVENTION: and Amplification of the Beta-glucosidase Gene of
 NUMBER OF SEQUENCES: 4
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Genencor International, Inc.
 STREET: 925 Page Mill Road
 CITY: Palo Alto
 STATE: California
 COUNTRY: U.S.A.
 ZIP: 94304-1013

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
 OPERATING SYSTEM: IBM PC compatible
 SOFTWARE: PC-DOS/MS-DOS
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/463,461
 FILING DATE: 05-JUN-1995
 CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Christopher L. Stone
 REGISTRATION NUMBER: 35,696
 REFERENCE/DOCKET NUMBER: GC78D4
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 415-846-7555
 TELEFAX: 415-845-6504

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:
 LENGTH: 3033 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 HYPOTHEtical: NO
 ANTI-SENSE: NO
 ORIGINAL SOURCE:
 ORGANISM: Trichoderma reesei
 FEATURE:
 NAME/KEY: CDS
 LOCATION: join(311..375, 446..2205, 2270..2679)
 FEATURE:
 NAME/KEY: Intron
 LOCATION: 376..445
 FEATURE:
 NAME/KEY: Intron
 LOCATION: 2206..2269
 US-08-463-461-1

Query Match

Best Local Similarity 2.8%; Score 83.8; DB 3; Length 3033;
 Matches 264; Conservative 0; Mismatches 257; Indels 6; Gaps 2;

QY 552 TCACATTCAACCAACTCTGCTCGAAGAGGCGAGTAAAGTATGGGCAAGAGCCATC 611
 DB 720 TCGACGTGGGATGTCATTTGATCCCGAAGCGAGACATTCATCGTGAAGAGTGAAG 779
 QY 612 GCTAAGAGTGGCATGTGATCTCGGCC--GACTATCAACATGCAACGCTCCCTCTC 668
 DB 780 GCTCGGGGATTCATGTCATCTGCTGCTGGGCGCGTGGAAAGACTCCGAG 839
 QY 669 GGTGACGTGGCTTCGAGTGGATGGTGGAGATCCCTTCCTGGCGGCTTGGAGCTCG 728
 DB 840 GGGGTGCGCACTGGAGGGCTTCGCTGTCATTCATTCATTCAGGCGCATTCGATGGGT 899

QY 729 GCTTCATCCGCGCATTCGAGACACTGAGAGTGCAGAGCTACGATCAGACCTTTTGTGC 788
DB 900 CAAACCATCAACGCGCATTCAGTCGTAGGCGTGCAGGGCAGCAGAGACTATATCCTC 959
QY 789 AATGATCAGAGAGACAGCGGCATGATGTCAGAGCATGTCACGAGGGGCTCTCCGT 848
DB 960 AACGAGCAGAGAGCTCAATCGAGAAACCATTTGAGCAACCCAGATGACCGAACTCTCCAT 1019
QY 849 GAAATCTACGACATCCCGTTCAGATTCGTGCGAGACTCCAGCGGGTGCCTTCATG 908
DB 1020 GAGCTGTATACTTGGCCATTTGCCGACGCGGTTCC--AGGCCAATGTCCCTTCTGTATG 1076
QY 909 ACGGCGTACATGCGATCATGAGGCGTGTGTCGACGAGAGAACCTTAATATCTTGATGG 968
DB 1077 TCGCTGTACACAGGTCAATACCACTGCGGCTGCGAGATCACTACACGCTGCAGACT 1136
QY 969 ATGCTTGAAGAGATGCGGTGGGATGGCCATATCATGAGCGACTGTACGGCACAATAC 1028
DB 1137 GTGCTGAAGAGACCACTGGGGTTCCAGGCTATGTATGACGAGCTGGAACGACACAGCAC 1196
QY 1029 AGTACACAGAGCCGTTGTGCGAGGCCCTCGACCTCGAGATGCCCG 1075
DB 1197 ACGACTGTCCAAAGCGCGAATTCGTGGCTTGACATGTCAATGCCCTGG 1243

Search completed: May 3, 2003, 22:13:50
Job time : 3750.91 secs

